p-n HETEROJUNCTIONS OF INDIUM ANTIMONIDE ON ELECTROPOLISHED AND CHEMICALLY POLISHED SILICON

by

RAMASWAMI SUBRAMANIAN

B. Sc., University of Madras, 1956

A MASTER'S THESIS

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Physics

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1966

Approved by:

Major Professor

TABLE OF CONTENTS

INTRODUCTION

A heterojunction is essentially a junction between two dissimilar semiconductors with different energy gaps. This junction may be p-n, p-p, n-p, or n-n in character. Junctions in a homogeneous semiconductor have been studied extensively and well understood, [1-3] whereas heterojunctions have received comparatively little attention. This is principally due to the difficulty of preparing well behaved junctions. However, because of the advances made recently in the growth of epitaxial films of Groups III-V compounds from vapor phase, preparation and properties of several heterojunction systems have appeared in recent literature. [6-9, 12, 13]

Indium antimonide (InSb) is a Group III-V compound and a semiconductor with an energy gap of 0.17 eV between conduction and valence bonds and a lattice constant of 6.49Å. Silicon (Si) is a Group IV semiconductor with an energy gap of 1.08 eV and a lattice constant of 5.40Å. It was desired to investigate a heterojunction that would result when these two materials are brought together. A vacuum evaporation technique [20] was used to deposit a thin film of InSb epitaxially on silicon substrates, thus forming a heterojunction between InSb and Si.

In order to deposit a highly perfect epitaxial film a substrate with an extremely smooth, flat and polished surface is required. The methods employed for this purpose are mechanical polishing combined with a short chemical etch, chemical polishing and the more recent electropolishing. [14, 15] The electropolishing technique was chosen for the preparation of the silicon substrates and an electropolishing mechanism was constructed.

In addition a chemical polishing technique was used in the preparation of other substrates. The epitaxy of InSb on Si was studied and the electrical characteristic of a p-n InSb-Si heterojunction were obtained and interpreted, by comparison with the results of Anderson for a p-n Ge-GaAs heterojunction.

THEORY

p-n Junctions and Heterojunctions

Theory of p-n junctions is reviewed here briefly, at first, and an extension of the theory is made to explain the behavior of heterojunctions on the basis of an energy band diagram as deduced by Anderson.[6,8] The following is a list of symbols and their meaning as used in the theory presented below.

D_p : Diffusion constant for holes

D_n : Diffusion constant for electrons

$\left.\begin{matrix} e \\ exp \end{matrix}\right\}$: Exponential function

E_{g1} , E_{g2} : Energy gaps in semiconductors 1 and 2 respectively

ΔE_g : Difference between the energy gaps

ΔE_c : Amount of discontinuity between the conduction band edges

j_p : Hole drift current density

j_n : Electron drift current density

J_p : Junction hole current density

J_n : Junction electron current density

J_o : The net dc current through a junction under applied forward bias

k : Boltzmann's constant : $1/11600$ eV/^{o}K

L_p : Diffusion length for holes : $\sqrt{D_p \tau_p}$

L_n : Diffusion length for electrons : $\sqrt{D_n \tau_n}$

n : Free electron concentration : number/cm^3

n_i : Intrinsic carrier concentration (electrons or holes)

n_p : Electron (minority carrier) concentration in p-type semiconductor

n_n : Electron (majority carrier) concentration in n-type semiconductor

N_D : Concentration of the donor impurity atoms

N_A : Concentration of the acceptor impurity atoms

p : Free hole concentration

p_n : Hole (minority carrier) concentration in n-type semiconductor

p_p : Hole (majority carrier) concentration in p-type semiconductor

q : Absolute magnitude of electronic charge : 1.6×10^{-19} coul.

T : Room temperature in degrees Kelvin : 300^{o}K (approximately)

v : Applied forward bias voltage

V_D : Equilibrium contact potential or built in voltage

V_{D1} , V_{D2} : Built in voltages in semiconductors 1 and 2 respectively

x : Coordinate of position inside p or n region in a semiconductor

x_p , x_n : Coordinates of position of the edges of the transition region in a p-n junction.

ϕ_p : Fermi energy level in terms of the potential energy in a p-type semiconductor.

ϕ_n : Fermi potential in an n-type semiconductor

μ_p : Hole mobility

μ_n : Electron mobility

τ_p : Mean life time of holes

τ_n : Mean life time of electrons

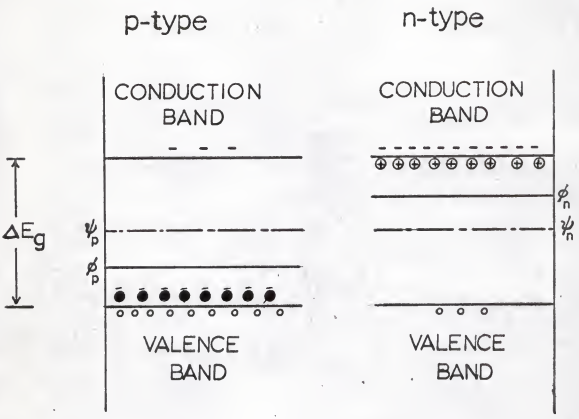
A homogeneous p-n junction or a homojunction is formed when a semiconductor doped with a p-type impurity is brought into contact with a semiconductor of the same material doped with an n-type impurity. The action of a p-n junction can be analyzed with the help of energy band diagrams. Figure 1 indicates the energy bands of two isolated p and n regions of the same semiconductor. The n-type impurity is usually from Group V and, when it displaces a Ge atom in the lattice, has a loosely attached electron. At room temperature these electrons cross over to the empty states in the conduction band and thus leave behind an ionized impurity atom. Thus this type of impurity is called a donor and the donor states when unoccupied give rise to electrons in the conduction band, the majority carriers in an n-type semiconductor.

In an intrinsic n-type semiconductor there is normally an equal number of electrons in the conduction band and holes in the valence band. Doping with an n-type impurity raises the concentration of electrons above the intrinsic level n_i and lowers the concentration of holes below the intrinsic level. Carrier concentrations for an n-type semiconductor can be derived in terms of the departure of the Fermi level from its intrinsic value

$$n_n = n_i \exp \left\{ q(\psi_n - \phi_n) /kT \right\} \qquad (1)$$

$$p_n = n_i \exp \left\{ q(\phi_n - \psi_n) /kT \right\} \qquad (2)$$

In the above equation, the potential ϕ and its corresponding energy $W = q\phi$ are referred to as Fermi level. The Fermi level is defined as that level which, if it were allowed, would have the probability of occupation for an electron of one-half. Similar arguments apply to a p-type semiconductor.



●: ACCEPTORS, ⊕: DONORS, -: ELECTRONS, ∘: HOLES, ψ: ELECTROSTATIC POTENTIAL AND ϕ: FERMI LEVEL.

Fig. 1. Energy bands and carriers in isolated p and n regions of a semiconductor.

Holes are the majority carriers in p-type semiconductors; they are caused by the vacancies created by electrons when they jump over from the valence band to fill the acceptor states created by acceptor atoms. Similar equations relating Fermi level and carrier concentrations in a p-type semiconductor are:

$$p_p = n_i \exp \left\{ q \, (\phi_p - \psi_p)/kT \right\} \tag{3}$$

$$n_p = n_i \exp \left\{ q \, (\psi_p - \phi_p)/kT \right\} \tag{4}$$

It should be noted that electrons are minority carriers in a p-type semiconductor.

When two such regions are brought into intimate contact, at thermal equilibrium, a contact potential is developed. The action of the p-n junction can be understood by considering the mechanism of current flow in a p-n junction.

The two processes by which current flow occur in a p-n junction are drift and diffusion. The carriers are free to move randomly through the crystal lattice. In moving through the crystal structure they encounter various retarding forces such as collision with impurity atoms, displaced lattice atoms, and other mobile carriers. However, when subject to a uniform external field, the random motion is averaged out and the carriers attain limiting drift velocity. Charge is thus transported by the drift of carriers and the current is called drift current. The equation for drift current density for holes in a p-type semiconductor in terms of its mobility, concentration p at a point where the electric field intensity is E units is given by

$$j_p = q p \mu_p E \tag{5}$$

A similar equation for electrons in an n-type semiconductor will be

$$J_n = q n \mu_n E . \tag{6}$$

Whenever there is a gradient in carrier density the process of diffusion takes place. By this process carriers move away from regions of high density towards those of lower density. Charge is transported thus by diffusion causing a diffusion current. At thermal equilibrium when p and n regions are brought together* diffusion currents are set up and holes in valence band from the p side diffuse into the lower density n region. Similarly electrons in the conduction band from the n side diffuse into lower density p region. An accumulation of charge takes place in the p and n regions with n becoming positive with respect to p. This sets up an electric field causing an opposing current due to drift. The system attains equilibrium when a potential equal to the contact potential is attained at which both drift and diffusion currents balance each other. This situation is illustrated in Fig. 2.

The total current across the junction for any one type of carrier is the algebraic sum of drift and diffusion components as shown by this equation for holes:

$$J_p = \underbrace{q \mu_p p E}_{\substack{\text{Drift} \\ \text{Component}}} - \underbrace{q D_p \nabla p}_{\substack{\text{Diffusion} \\ \text{Component}}} \tag{7}$$

Approximating by one dimensional analysis,

---

*The theory of p-n junctions in a homogeneous semiconductors presented here in pages 7—14 , is after Middlebrook. (1)

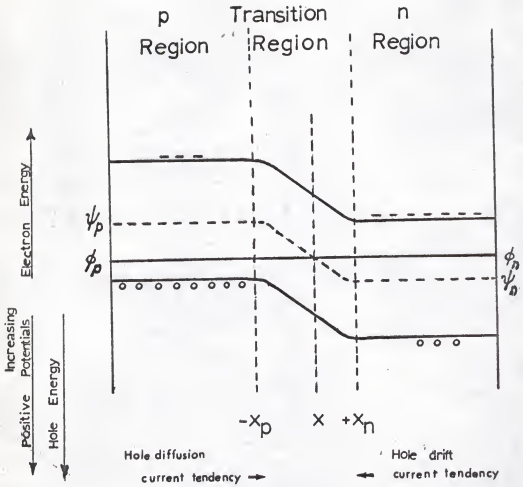


Fig. 2. Energy band diagram of a p-n junction at thermal equilibrium.

$$E \doteq -\frac{d\psi}{dx} \tag{8}$$

$$\nabla p \doteq \frac{dp}{dx} \tag{9}$$

Substitution in equation (7) gives

$$J_p = -q\mu_p p \frac{d\psi}{dx} - qD_p \frac{dp}{dx} . \tag{10}$$

However, in equilibrium conditions, there can be no net transfer of charge from any one region to the other. The balance between density gradient force and field force must therefore be obtained at every point across the junction. Integrating this difference over the transition region and setting it equal to zero, (that is setting effectively, $J_p = 0$), one obtains, after substitution of appropriate boundary conditions:

$$q\mu_p \int_{\psi_p}^{\psi_n} d\psi + qD_p \int_{p_p}^{p_n} \frac{dp}{p} = 0 \tag{11}$$

which leads to

$$p_p = p_n \exp\left\{ (\psi_n - \psi_p)\, \mu_p / D_p \right\} . \tag{12}$$

Which is a relation between hole density at the edge of the transition on the p side to that at the edge of the transition region on the n side. However, the equilibrium concentrations of p_n and p_p are given by equations (2) and (3). Substitution yields the Einstein relation for holes:

$$D_p = \frac{kT}{q} \mu_p \; ; \tag{13}$$

A similar relation for electrons given by:

$$D_n = \frac{kT}{q}\mu_n \quad . \tag{14}$$

A p-n junction under applied dc potential is illustrated in Fig. 3. The forward bias making p positive with respect to n, effectively shifts downward the energy bands in the p region. This shift in turn reduces the effective barrier for holes. Minority carrier injection occurs and similarly the potential barrier for electrons is reduced and more electrons spill over to the p side from the n side. The procedure employed for solving the equilibrium case is readily extended to the junction under an applied potential. The equation (10) for the net current transported by holes across the junction may be rearranged and integrated with the boundary conditions as:

$$q\mu_p \int_{\psi(-x_p)}^{\psi(+x_n)} d\psi \;+\; qD_p \int_{p(-x_n)}^{p(+x_n)} \left(1 + \frac{J_p}{qD_p}\frac{dx}{dp}\right)\frac{dp}{p} = 0 \tag{15}$$

In the second integrand the term involving J_p may be written as $\dfrac{J_p}{qD_p \dfrac{dp}{dx}}$.

The denominator is recognized as the diffusion current density and, calculating for practical conditions in a semiconductor, one obtains a very large value for the diffusion current density of the order of 25000 A/cm^2. Such large current densities, however, do not occur in a p-n junction, whether in an equilibrium or under an applied potential, since this large diffusion current flow is matched by a drift current of the same order of magnitude in the opposite direction. Thus the second term in the integrand may be neglected and the equation for the junction under applied potential

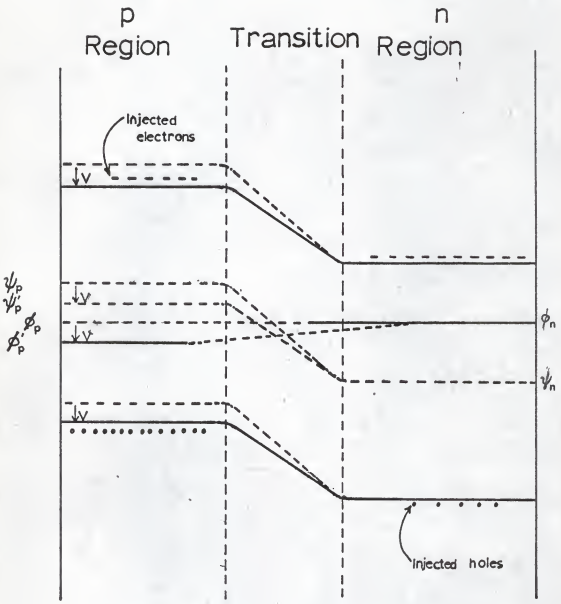


Fig. 3. Energy band diagram of a p-n junction under applied forward dc potential.

reduces to:

$$q\mu_p \int_{\psi_p+v}^{\psi_n} d\psi \quad + \quad qD_p \int_{p_p}^{p_{eo}} \frac{dp}{p} \quad = \quad 0 \tag{16}$$

with the appropriate boundary conditions. Integration yields

$$p_p = p_{eo} \exp \left\{ (\psi_n - \psi_p - v) \mu_p / D_p \right\} \tag{17}$$

where p_{eo} is the minority hole density just inside the n region. The majority carrier densities away from the transition region remain effectively at the equilibrium values on either side whereas the minority carrier densities are increased due to injection. Eliminating p_p using equation (12) gives

$$p_{eo} = p_n \exp(\mu_p / D_p) \tag{18}$$

which is the injected hole density just inside the n region in terms of the junction voltage.

Electrons and holes in a p-n junction must satisfy continuity equation which may be stated physically as:

$$\left[\text{Time rate of increase of holes}\right] = \left[\text{rate of thermal generation of holes}\right] - \left[\text{rate of recombination of holes}\right] - \left[\text{divergence of hole flow}\right].$$

Solving the continuity equation, assuming one dimensional flow and further assuming that injected hole current is much larger compared to minority

drift current, one obtains the expression for injected hole density as a function of the position in the n region:

$$p_o - p_n = p_n \left(e^{qv/kT} - 1\right) \exp\left\{(x_n - x)/L_p\right\} \tag{19}$$

Subscript o stands for dc conditions. The injected hole current density is readily obtained after differentiating p_o:

$$J_{po} = -qD_p \frac{\partial p_o}{\partial x} = \frac{qD_p p_n}{L_p}\left(e^{qv/kT} - 1\right) \exp\left\{(x_n - x)/Lp\right\} \tag{20}$$

where x_n gives the position of the edge of the transition region. At $x = x_n$:

$$J_{po}(x_n) = \frac{qD_p p_n}{L_p}\left(e^{qv/kT} - 1\right) \tag{21}$$

The net current through the junction is the sum of the injected hole and electron current given by:

$$J_o = \left[\frac{qD_p p_n}{L_p} + \frac{qD_n N_p}{L_n}\right]\left(e^{qv/kT} - 1\right) \tag{22}$$

The minority concentrations p_n and n_p in n and p regions may be expressed in terms of the respective majority carrier concentrations and equilibrium contact or built-in potential using equations (12 and 13), by the relations

$$p_n = p_p \exp\left\{-qV_D/kT\right\} \tag{23}$$

$$n_p = n_n \exp\left\{-qV_D/kT\right\} \tag{24}$$

where $V_D = \psi_n - \psi_p$.

However, the majority concentrations are approximately equal to the impurity concentrations at room temperature, given by the relations

$$n_n \doteq N_D \tag{25}$$

$$p_p \doteq N_A \tag{26}$$

Incorporating the Eq. (23,24,25 and 26) in Eq. (22) gives the rectifier characteristic.

$$J_o = \left[\frac{qD_p N_A}{L_p} + \frac{qD_n N_D}{L_n} \right] \exp(-qV_D/kT)\,(e^{qV/kT} - 1) \tag{27}$$

$$= \left[qN_A \frac{D_p}{p} + qN_D \frac{D_n}{n} \right] \exp(-qV_D/kT) \left\{ e^{qV/kT} - 1 \right\} \tag{28}$$

using the expressions for diffusion lengths $L_p = \sqrt{D_p \tau_p}$ and $L_n = \sqrt{D_n \tau_n}$. The rectifier characteristic is usually expressed in the form

$$J_o = J_s\,(e^{qV/kT} - 1) \tag{29}$$

where

$$J_s = \left\{ qN_A \sqrt{\frac{D_p}{\tau_p}} + qN_D \sqrt{\frac{D_n}{\tau_n}} \right\} \exp(-qV_D/kT) \tag{30}$$

is the saturation current under applied reverse bias. Figure (4) illustrates the rectifier characteristic of a p-n junction.

Theoretical study of a heterojunction is by no means as simple. Abrupt heterojunctions are epitaxial semiconductor-semiconductor junctions formed between dissimilar materials with different work functions, electron

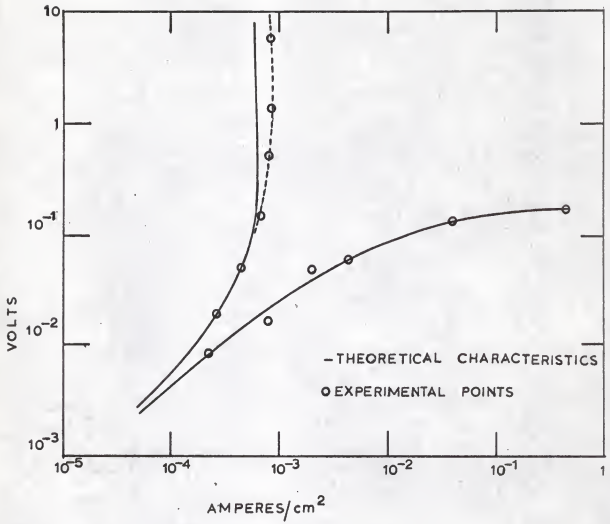
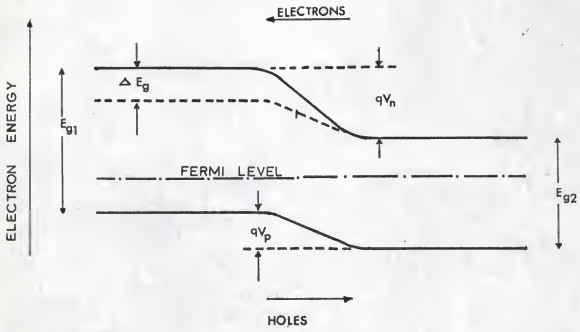


Fig. 4. Rectifier characteristic of a p-n junction (after Middlebrook (1)).

affinities, and forbidden energy gaps. The junction formed is characterized as p-n, p-p, n-n, or n-p if the respective semiconductors have appropriate dopants. Due to the epitaxial nature of the junction, a good lattice match between the pair of semiconductors is desired, so that there will be negligible strain at the interface and few dislocations. Few semiconductor pairs can be found with the closeness of lattice constants desirable. Oldham and Milnes [10] have shown that due to the lattice mismatch, the unavoidable interfacial dislocations in the heterojunctions of present interest are likely to produce allowed energy states in the energy gap affecting the device characteristics. Thus from a practical standpoint, an exact analytical solution to the problem of heterojunction is quite difficult. However, there have been a few theories proposed in literature [6-9] which will be reviewed briefly here.

First, Shockley [5], and Kroemer [4] suggested using a wide-gap semiconductor as an emitter for a transistor which has a base and collector of a narrow gap semiconductor, to obtain a high injection efficiency. Fig. (5) shows energy band model of a p-n heterojunction as proposed by Kroemer. The current gain β of a transistor in grounded emitter operation is given by the expression $\beta = \frac{\alpha}{1-\alpha}$ where α is the short-circuit forward-current transfer ratio. The static value of α is given by the ratio of collector current to emitter current and is thus also termed the current amplification factor of the emitter. α, which has a value close to unity, is proportional to the emitter efficiency which is defined as that fraction of total current that is minority carrier injection current. For a p-n-p transistor this may be expressed as

$$\alpha \propto \gamma = \frac{J_p}{J_n + J_p} \tag{31}$$



$\Delta E_g = E_{g1} - E_{g2}$

V_p = POTENTIAL BARRIER FOR HOLES

V_n = POTENTIAL BARRIER FOR ELECTRONS

Fig. 5. Energy band structure of a p-n heterojunction at equilibrium according to Kroemer (4).

For a high injection efficiency $1-\gamma$ should be small or be very close to unity.

$$(1-\gamma) \doteq \frac{1-\gamma}{\gamma} = \frac{J_n}{J_p} \qquad (32)$$

$1-\gamma$ is called the injection deficit. Kroemer compared the injection deficits in a p-n homojunction and a p-n heterojunction and showed that, assuming the effective masses of holes to be identical, the injection deficit of a heterojunction is reduced by a factor of $\exp(-\Delta Eg/kT)$, from that of a homojunction. This factor can be shown to be nearly 1:3000 for a $\Delta Eg = 0.2$ eV. The response to the signal frequency of α in a homojunction is normally such that it has a constant value region and then it begins to decrease rapidly. The frequency at which the value of α is 3 db below the dc value is termed the α fall-off. α fall-off significantly depends on the minority carrier storage and thus on injection deficit. Since the injection deficit in a heterojunction is 3000 times smaller, α fall-off would occur in a heterojunction if the emitter current were 3000 times that of the homojunction. Thus for normal operating currents, α fall-off in the case of a heterojunction is practically non-existant. Alternatively in small signal transistors where it is not of prime importance to minimize the alpha fall-off, one can use this effect to diminish the effect of the doping in the emitter by the same amount, and still have an unchanged fall-off characteristic with much reduced emitter transitional capacitance. Attempts to fabricate such a transistor have not been successful until recently.

Fabrication techniques for several heterodiodes however, have been developed in recent years and their characteristics were reported. (6-9, 13)

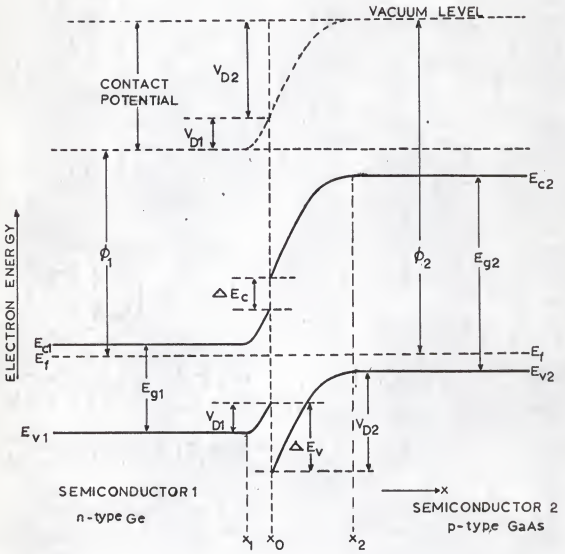
The theory that has so far best explained the experimental results has been proposed by Anderson (8).

In the energy band diagram for heterojunctions Figs. (6a, b, and c) that Anderson proposed, the conduction and valence band edges are discontinuous at the interface. The discontinuities arise from the fact that there is an abrupt change in the semiconductor energy gap and electron affinity at the junction interface. Though theoretical justification of such discontinuities in the band edges at the interface is difficult, the Anderson model has been considered a good approximation for analysis of abrupt structures like heterojunctions.

On the basis of the above model, the electrical properties of the junction can be predicted. It is seen that the current in a heterojunction is almost entirely due to one type of carrier, either holes or electrons, since the potential barriers to the two types of carriers have different magnitudes. As an example a n-p heterojunction is considered in Fig. (6a). The barrier for holes (V_{D2}) is much smaller than the barrier for electrons ($V_{D2} + V_{D1} + \Delta E_c$), where V_{D1} and V_{D2} are the built in voltages respectively in dissimilar n- and p-type materials. Thus current flow would be expected to be mainly by holes. Generalization of Eq. (28) derived for a homojunction, for the case of a n-p heterojunction yields:

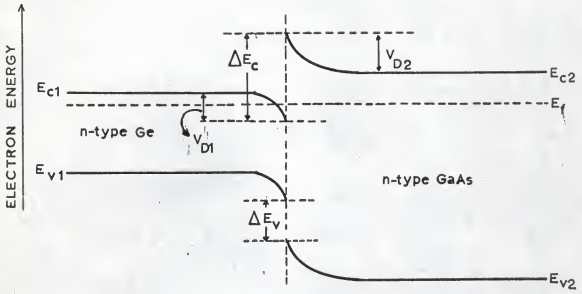
$$J_{n-p} = B \exp(-qV_{D2}/kT) \left\{ e^{qv/kT} - 1 \right\} \qquad (33)$$

where $B = qN_{A2}\sqrt{\frac{D_p}{\tau_p}}$ and V_{D2} is the partial built-in voltage or the electrostatic potential supported by semiconductor 2. The equation for a p-n junction would be similar.

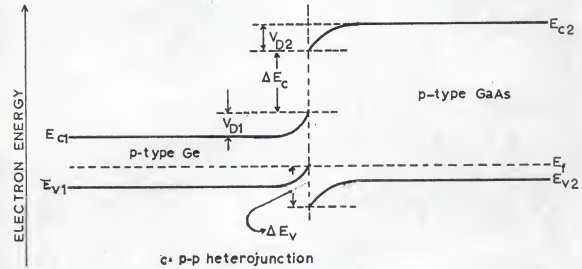


E_f: FERMI LEVEL, E_c = CONDUCTION BAND EDGE , E_v: VALENCE BAND EDGE, ϕ: WORK FUNCTION, V_D = BUILT-IN VOLTAGE, AND E_g: ENERGY GAP. SUBSCRIPTS 1 AND 2 DENOTE RESPECTIVE SEMICONDUCTORS.

Fig. 6a. Anderson model of a n-p heterojunction at equilibrium (after Anderson (8)).



b. n-n heterojunction



c. p-p heterojunction

Fig. 6b and c. Energy band models of n-n and p-p heterojunctions at equilibrium (after Anderson (8)).

Anderson predicted the current-voltage characteristics of n-n and p-p junctions using diode emission theory and found:

$$J_{n-n} = B' \exp(-qV'_{D2}/kT) \left\{ e^{qv/kT} - 1 \right\} \qquad (34)$$

where $B' = qN_{D2} \left(\frac{kT}{2\pi k}\right)^{1/2}$ (35)

Further, in order to explain the reverse current-voltage characteristic, potential barriers (V_{D2}, V'_{D2}) in Eqs. (33, 34) were assumed to vary with the applied reverse voltage and the variation was attributed to image effects and quantum mechanical tunnelling. Anderson found good agreement between theoretical and experimental reverse current-voltage curves when the above mentioned effects were taken into account.

On the basis of the preceding theory an attempt is made to give an energy band diagram and predict the current-voltage characteristics for a p-n InSb-Si heterojunction, with which this work is primarily concerned. Values of the lattice constant and energy gap of 5.40 Å, 1.08 eV for Si [9] and 6.49 Å, 0.17 eV for InSb [22] were used. These values were used in the energy band diagram for the p-n InSb-Si heterojunction, as shown in Fig. (7).

It is seen that the barrier for electrons V_{D2} is much smaller than for holes $(V_{D2} + V_{D1} + \nabla E_V)$. The predominant carriers are then electrons. A contact potential is established at thermal equilibrium with Si positive with respect to InSb. Hence, for a forward bias InSb should be made positive with respect to Si. It is expected that current-voltage characteristics, will be given by the equation:

$$J_{p-n} = qN_D \sqrt{\frac{D_n}{\tau_n}} \exp(-qV_{D2}/kT) \left\{ e^{qv/kT} - 1 \right\} \qquad (36)$$

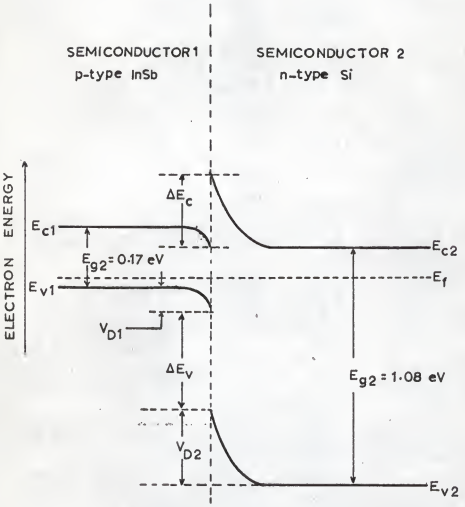


Fig. 7. Energy band diagram of a p-n InSb-Si heterojunction at equilibrium.

where N_D is the donor impurity concentration in Si which is inversely proportional to the resistivity of Si and V_{D2} is the partial built-in voltage in silicon of the InSb-Si heterojunction. Thus the current in this case will depend significantly on the resistivity of silicon and the built-in voltage supported by silicon.

It is expected that the forward current should vary approximately exponentially and the reverse current should show a soft breakdown in a manner similar to the current-voltage curves. In the above theory presented here no attempt is made to draw a theoretical curve, but the expected results of a general nature are summarized.

A more rigorous treatment of the p-n heterojunction is found in a recent publication by Pearlman Feucht. (11) A n-n or p-p heterojunction has entirely electrons (in the case of n-n) or holes (in the case of p-p) as carriers and thus there are negligible minority carrier storage effects. Thus, very fast switching has been realized with n-n heterodiodes. (9) Similarly n-p and p-n heterojunctions possess considerably less storage effects since conduction is mostly by one type of carrier, the barrier to the other being very large. All the physical processes in a heterojunction are not completely understood. However heterojunctions are being considered for several possible uses. One recent proposal has been the use of a pair of heterojunction injectors for laser action. (24).

Anodic Dissolution and Electropolishing of Silicon

Most of the heterojunctions developed to date have been made by epitaxial vapor deposition method. The active regions of such devices are so close to

the surface of the semiconductor that minor imperfections in the surface structure of the substrate on which the epitaxial film is formed, can significantly affect the physical and electrical properties of the device. In order to realize nearly perfect epitaxial films and thereby good heterojunctions, substrates, which posses extremely flat surfaces, free of all work damage and etch pits, are required.

Conventional methods of preparation of a substrate surface by mechanical polishing followed by a short chemical etch are far from satisfactory in producing damage free smooth and flat surfaces. This problem has been overcome in the development of an electropolishing technique for silicon and germanium by Sullivan et al. (14) This method for Si has been further investigated by Baker and Tillman. (15) Successful electropolishing of silicon, of various orders of resistivities, has been reported using the technique reviewed briefly here.

The principle, which involves anodic dissolution of silicon, is discussed thoroughly in a comprehensive review of electrolyte-semiconductor behavior by Turner. (16) Acidic solutions containing fluoride ions have been found to be suitable electrolytes for silicon dissolution. Sullivan used dilute NH_4HF_2 (ammonium bifluoride) where as Baker and Tillman used dilute solutions HF (hydrofluoric acid). The mechanism of electropolishing as proposed by Turner, involves effectively an electrochemical process as given by the equations:

$$\begin{array}{l} Si + 2H_2O + Ze^+ \longrightarrow SiO_2 + 4H^+ + (4 - Z)e^- \\ \qquad\qquad\qquad\quad \Big\downarrow 6HF \\ \qquad\qquad\qquad\quad H_2SiF_6 + 2H_2O \end{array} \qquad (37)$$

the e^+ being holes in the valence band and the e^- are electrons which are the carriers. Electrons are readily available from broken semiconductor bonds whereas the holes must be supplied. This point will be discussed later. The overall process may be explained as follows: In electropolishing, a thin film is present on the surface which is postulated to be a thin film of SiO_2. During anodic dissolution the surface layer of silicon contains approximately a mono layer of the oxide. The oxide of silicon is not water-soluble and it remains on the surface until sufficient HF reaches the surface by mass transfer to form water-soluble fluoride complex. Four hydrogen ions are generated for each Si atom in the dissolution process, as has been verified by the presence of gassing during electropolishing. Further it is seen that, in the process of dissolution of Si atoms, holes are removed. Thus a copious supply of holes is required for the process to go on continuously. In the case of n-type silicon, this poses a serious problem especially for silicon with a resistivity greater than 0.01 ohm-cm. Thus an external supply of holes by injection are needed for uniform electropolishing. Sullivan (14) employed light injection as a means of generating the holes needed for the dissolution.

Turner and others (16) found that an intensive stirring action was necessary to obtain the special effects during anodic dissolution. Sullivan accomplished the intensive stirring by a system of rotating electrodes with a thin film of electrolyte between them. The anode-cathode spacing was kept small to obtain a high degree of polish and flatness. The potential-current density curve for the electrolyte as found by Baker and Tillman is given in Fig. (8). A similar characteristic has been reported by Sullivan (14) and others (16) before. As seen in the figure, at low voltages a thick film formation takes place. A critical current density was found necessary before

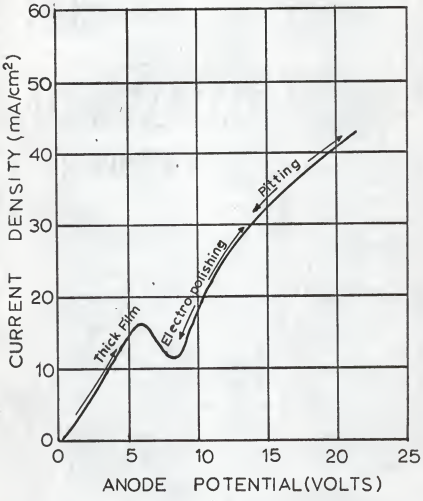


Fig. 8. Typical potential-current density curve for the electrolytic cell of the electropolishing mechanism (after Baker and Tillman (15)).

electropolishing starts. At high current densities pitting is found to result and can be avoided if the current is maintained at a value slightly above the critical value in the electropolishing region. Further for n-type Si of resistivity above 0.1 ohm-cm Sullivan found that the concentration of the electrolyte must be lowered, in order to produce a good polish.

Thus by proper choice of the electrolytic cell voltage, current and the concentration of the electrolyte, it is possible to obtain good results. Sullivan and Baker have reported extremely good results with electropolishing of both types of Si of various resistivities in the range of 0.001 to 300 ohm-cm.

EXPERIMENTAL TECHNIQUES

Technique of Electropolishing

A substrate of Si with a perfectly smooth and flat surface is needed to form a good heterojunction. To fulfill this requirement an electropolishing device was constructed, applying the principles outlined in the theory above, and modeled after the one developed by Sullivan. (14)

Figure (9) gives the schematic of the electropolishing equipment. The essential components are the large motor driven cathode and a small non-conducting disc carrying the samples of silicon which form the anodes. The cathode was tilt mounted as shown with its plane making an angle of 15 degrees to the vertical. This type of mounting together with gravity facilitated the close positioning of the anodes required. The cathode was constructed from

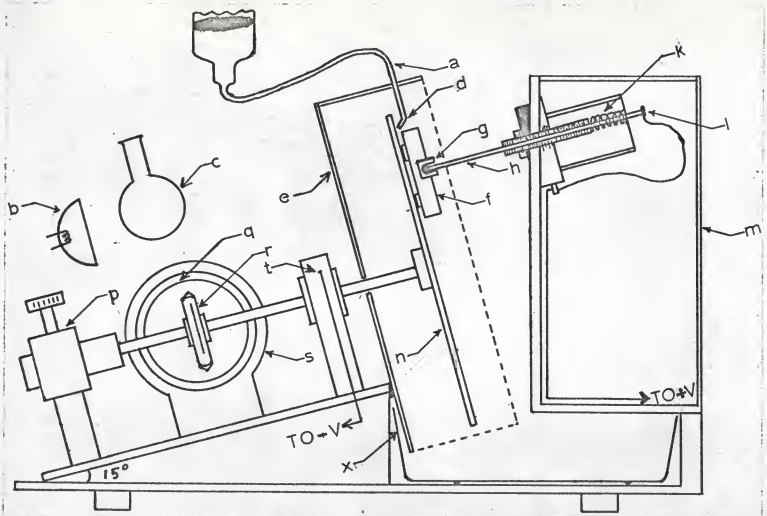


Fig. 9. Schematic of the electropolishing mechanism. (a) Etchant flow line. (b) lamp. (c) Cooling flask. (d) Nozzle. (e) Plexiglass frame. (f) Anode block. (g) Ball and socket joint. (h) Ball ended shaft. (k) Spring loading. (l) Anode terminal. (m) Support frame. (n) Cathode. (p) Rack and pinion. (q) Friction drum. (r) Friction wheel. (s) Motor. (p,q,r and s comprise variable speed drive). (t) Cathode terminal. (x) Poly ethylene drain.

rolled steel disc about 8" in diameter with perforations all throughout its area in order to make it semitransparent when rotated at a speed of about 100 r.p.m. or less. This semitransparency of the cathode facilitated injection of holes on the surface of the anode samples, by allowing the focusing of a light beam on the anode surfaces through the cathode.

For such an application, a high degree of flatness of the cathode was necessary. The cathode surface was carefully ground flat and it was mounted to the shaft so that its plane was perpendicular to the axis of the shaft. The shaft was axially coupled to the driving motor through a variable speed friction wheel drive. (see Fig. 9) To make the cathode withstand the corrosive action of the electrolyte it was gold plated. The plating should be thick enough to withstand the intensive rubbing action of the anodes as well. The plating, done by Signumd Cohn Co., had a coat of 0.0001" Cu and then a coat of 0.0001" Ni and finally was electroplated with gold in a heavy plating solution with about six millionths of gold on it.

The anode block was a bakelite disc 3" in diameter. The schematic of the block is shown in Fig. (10). The brass plug that held the silicon anodes was fitted into the plexiglass coaxial cylinder which, itself, was movable in and out of another coaxial cylinder pressed into the bakelite disc. The disc possessed three such arrangements symmetrically located. Such an arrangement was flexible enough so that if silicon samples of different thicknesses were to be polished they could all be mounted on the block so that their surfaces were all in the same plane. Set screws held the plugs and cylinders in position.

In order to keep a thin film of electrolyte it was necessary to hold the anodes very close and plane parallel to the cathode and yet sufficiently free

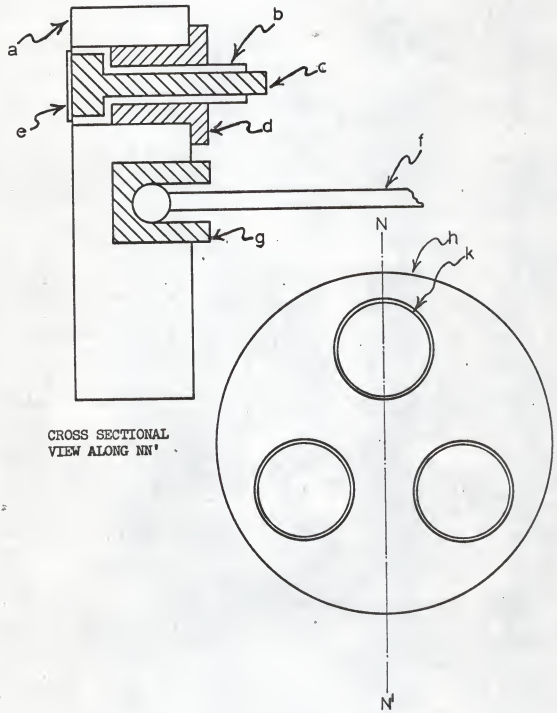


Fig. 10. Anode block assembly of the electropolishing mechanism. (a) Bakelite disc. (b) Moveable coaxial plexiglass cylinder. (c) Coaxial brass plug. (d) Coaxial plexiglass cylinder tight fitted. (e) Silicon sample to be electropolished. (f) Ball ended shaft. (g) Socket of the ball and socket joint. (h) Bakelite disc. (k) Silicon sample.

to rotate so that stirring action was accomplished. This was done by means of a ball and socket joint. The socket was tightfitted axially to the disc on the rear. The ball ended shaft on which it was positioned was spring loaded. The spring loading was specially designed so that it was flexible and the tension could be varied. The shaft with its spring loading was mounted to the frame facing the cathode by means of an adjustable screw arrangement, so that the length of the shaft carrying the ball extending beyond the frame could be adjusted. The purpose of the spring loading was to take up the extra tension caused by any departure from flatness of the cathode surface and the very slight wobbling which could not be totally avoided. Effectively the spring loading contributed to the uniform and smooth rotation of the disc. The main factors which contributed, however, are the tilted position of the cathode and the ball and socket joint, which together with gravity enabled the disc to ride smoothly on the electrolytic film on the surface of the cathode. This form of arrangement dispensed with very close mechanical tolerances.

The rotation of the disc was in effect due to the differential velocity developed by the viscous drag between the ends of the disc. The rotating electrodes were enclosed in a plexiglass frame which also partly collects and drains the electrolyte thrown from the periphery of the rotating cathode. Electrolyte was fed through a poly ethylene container. The rate of flow was controlled by means of a stop cock. An adjustable plexiglass frame at the top positioned the nozzle very close to the anode. The used electrolyte was collected in the poly ethylene trough at the bottom. Conductive lubricant containing colloidal graphite in the ball and socket joint assured proper electrical contact.

The silicon to be electropolished was cut in the form of thin slices of about 1 to 3 mm in thickness using a diamond cutting saw. Their diameter varied from 10 to 20 mm. After cutting, the slices were cleaned in trichloroethylene, followed by alcohol to remove grease and then rinsed in distilled water. The slices were then initially coarse ground on both sides with 40 micron particle size grinding powder on a glass lap. The edges were also ground smooth to remove any sharp corners acquired in cutting. The slices were then ultrasonically agitated in distilled water for about a minute to remove all adhereing particles. They were removed with a pair of tweezers from the ultrasonic bath and rinsed again in distilled water and then were soaked in hot HNO_3 for a few minutes. They were rinsed in distilled water after the soaking, and then were masked on one side with a thermosetting adhesive tape, on the side chosen for electropolishing.

Immersion in a bath of electroless nickel plating solution at a temperature of 95°C was followed, for about 2 minutes to give a good plating of nickel on the silicon slices needed for ohmic contact to the silicon. The electroless nickel plating solution [18] used was prepared with the following salts: nickel chloride ($NiCl_2 \cdot 6H_2O$), 30gm; sodium hypophosphite ($NaH_2PO_2 \cdot H_2O$), 20 gm; ammonium citrate $\{(NH_4)_2HC_6H_5O_7\}$, 65 gm; and, ammonium chloride (NH_4Cl), 50 gm. These salts were dissolved in a 1000 ml of distilled water and the bright green solution was filtered of impurities and stored in a sealed poly ethylene bottle. Before using the solution for nickel plating, ammonium hydroxide was added until the solution turned dark blue. The sample was immersed in the boiling plating solution for two minutes to obtain a sufficiently heavy plating for soldering.

In this electroless process, electron interchange takes place between two chemical species. The hypophosphite ion gives up electrons and becomes

phosphite ion. The nickel ion receives electrons and is deposited on the silicon. It was found in practice that a higher concentration of hypophosphite than was given in [18] was necessary to obtain Ni plating that would not peel off. The plated silicon slices were soldered on to brass plugs. It may be noted that, since the samples used were all of the high resistivity type, it was necessary to nickel plate the silicon and then solder in order to avoid the resistance spread in the silicon and make proper electrical contact.

While soldering the silicon slices on to the brass plugs care was essential to insure that there remained only a thin film solder between the brass plug and the silicon. The mounted slices were then slid carefully into the coaxial plexiglass cylinder and fixed in position with wet screws, so that silicon underside rested in flush with the plexiglass rim. The three plexiglass cylinders with the three silicon slices were then mounted on the anode disc. By placing the anode disc on a smooth plane glass plate, they were adjusted so that all the three silicon surfaces are in the same plane, and fixed in position by means of set screws. The anode disc was then mechanically ground and polished in the following steps prior to electropolishing.

This was done by lapping with Al_2O_3 abrasive powders of 17, 5 and 0.1 microns particle size on a glass plate using distilled water as a lubricant. After each operation of lapping with a particular size abrasive, the anode disc was cleaned ultrasonically in distilled water. The cleaned anode disc was dried using Kleenex and a thick coating of heavy solution of wax dissolved in trichloroethylene was used to seal exposed brass plug and also the cracks between plexiglass rim and the bakelite disc. This was then allowed to dry for one hour, before the final lapping operation was performed. Leads of copper wire were connected on the rear, to the socket at the center of the

disc by means of soldering. In order to observe the effect of electropolishing only two of the plugs were connected to the power supply.

To avoid direct electrical contact between anode and cathode a thin separator was used. After trying several kinds of fabrics, thin cloths of nylon or cotton about 75 to 100 microns in thickness were found to be best suited as a separator. The cotton was found to have more wet strength, and to withstand the rubbing action of the spinning anodes for a longer period of time. Illumination for light injection of holes was accomplished using a 250 watt aviation lamp focused into an intense beam. To cool and focus the beam, a flask of water was placed in front of the cathode-anode system. Cooling was required to prevent the heating of the wax sealant used on the anode disc.

The electrolyte used for electropolishing was made up of a solution of 0.5% NH_4HF_2 in 70% water and 30% glycerine. The glycerine served as a lubricant for smoother rotation of the anode disc and also to provide the viscous drag. A flow rate of approximately 30 ml/min. was found to be the optimum.

Before starting the electropolishing, the cathode surface was at first dampened with a solution of 30% glycerine and 70% water. The anode was positioned on the pivot of the ball ended shaft. The spring loaded shaft was also tilt mounted on the frame so that it was along the axis of the anode disc. Fine adjustments were made until there was very little tension on the anode disc so that it would spin freely when the cathode was rotated manually. Flow of the electrolyte was started, the light source was turned on and finally the dc voltage was applied, simultaneously starting the cathode drive motor. The device was allowed to run for about 45 minutes. Current through the electrolytic cell was measured by a millammeter. Voltage was directly read off the power supply. The operating voltage chosen depended on the resistivity of the sample and the concentration of the electrolyte, and was in the range

of 10-20 V. Current densities between 30 ma/cm^2 and 50 ma/cm^2 were obtained. A typical current voltage characteristic was obtained by varying the dc voltage. This is shown in Fig. (11).

After electropolishing the anode disc was detached and rinsed in distilled water. The brass plugs carrying the silicon were carefully unfastened from the disc. The silicon slices were removed from the plugs by careful heating. The silicon slices were then washed and rinsed clean in trichloroethylene to remove remanants of wax and glycerine. The electropolish obtained on n-type slices with a resistivity greater than 100 Ohm-cm was poor and considerable pitting was noted on the surface. Such slices were given a HF dip for about a minute and a half and rinsed thoroughly in distilled water followed by rinse in alcohol to remove moisture. The slices were then ready for transferring into the substrate holder in the vacuum system for depositing InSb film. The electropolished slice was also cleaned in trichloroethylene, rinsed in alcohol, and then transferred to the substrate holder in the vacuum system for the deposition of InSb. The substrate holder also contained a few other slices polished chemically. In order to keep the surface oxidation to a minimum, the substrate holder with the slices was transferred to the vacuum system immediately.

Techniques of Chemical Polishing

A chemical polishing method was also tried in the preparation of substrate surfaces for the deposition of InSb films. Chemical etching of semiconductors has been investigated by several workers. A good review of chemical etching was made by Irving [17] and also by Holmes [19]. Controlled

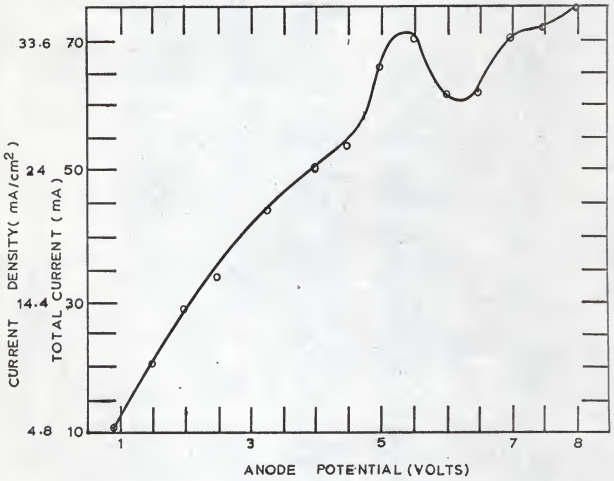


Fig. 11. Potential-current density curve for the electropolishing of silicon obtained with an electrolyte of 1% NH_4HF_2 dissolved in 30% glycerine and 70% water.

etchings of silicon in $HF-HNO_3$ systems was investigated by Klein and Stefan [25] who found that temperature control, stirring and relative concentrations of HF and HNO_3 are the factors influencing the etch rate and the degree of flatness of the etched sample.

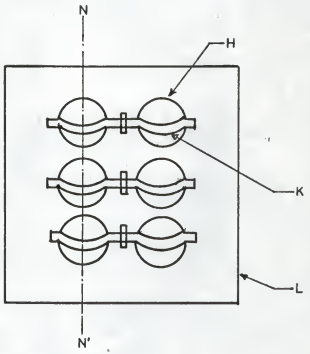
A chemical etch solution which would effectively polish the silicon and not preferentially etch it or reveal twins or dislocations etc., was desired. White etch, consisting of one part by volume of 48% hydroflouric acid to 3 parts by volume of concentrated reagent grade nitric acid was found to be the most suitable. The slices of silicon to be chemically polished were initially lapped on a glass plate with Al_2O_3 abrasives of particle size 40 microns, 17 microns, and 5 microns. After each lapping they were cleaned ultrasonically in distilled water before proceeding to the next step. After final lapping and rinsing the slices were cleaned in trichloroethylene, followed by rinsing in alcohol and distilled water. A special etching apparatus was constructed. It consisted of a small polyethylene trough with holes in its bottom. It was divided into four compartments and was attached to a long polyethylene stem for easy manipulation. A larger trough into which the smaller one could be immersed completely served as container of the etchant. This container was kept in a constant temperature bath of water at room temperature. The trough with the slices in the compartments was immersed in the etchant for 15 to 20 seconds, with constant agitation. Care was taken to see the silicon was always completely immersed. After 20 seconds, the reaction was stopped by diluting the etchant with a large amount of distilled water. After etching, the slices were rinsed in distilled water and finally in alcohol to remove all the moisture. As a result of chemical polishing, a very smooth and shiny surface was obtained. The slices were then transferred to the substrate holder inside the vacuum system for the deposition of InSb film.

Evaporation

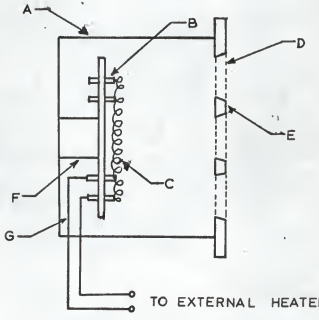
The recent development of techniques of epitaxial deposition has resulted in considerable advance in the preparation of heterojunctions. Most of the heterojunctions reported to date (6-9, 13) have been made by a vapor phase chemical reaction technique for the growth of an epitaxial film of Group III-V compound on the substrate which itself may be a Group III-V compound or Group IV semiconductor. Thus, most of the heterojunctions are three or more element systems such as: GaAs-Ge, GaAs-GaP, GaP-Ge, InP-GaAs, etc.. The usual technique of epitaxial deposition in these heterojunction systems mentioned above involved a chemical reaction of a Group III halide with a volatile Group V species either in a closed tube or in an open tube. In this present work, a vacuum evaporation method was used to develop an InSb-Si heterojunction. E. Brock Dale and An-Ti Chai [20] have developed the single source vacuum evaporation technique for InSb, used in this work. A tantalum crucible carrying a charge of InSb was used in conjunction with an Ultek high vacuum system for evaporation.

With a single source evaporation technique of this kind, the rates of evaporation of constituents of the InSb were properly controlled so that the deposited film was of the compound and not an uneven mixture of the two constituents.

Inside the bell jar of the vacuum system, two different substrate holder arrangements were provided, for mounting the substrates. One of the arrangements shown in Fig. (12) was a fixed, box shaped substrate holder, that carried exclusively the silicon substrates. This substrate holder and mask also contained a heating element to heat the substrates to a



SUBSTRATE MASK AND HOLDER



SECTION ON NN': HEATER

Fig. 12. Schematic diagram of the substrate mask and heater. (A) Aluminum casing. (B) Insulated metal posts. (C) Heating element. (D) Port. (E) Substrate mask. (F) Ceramic spacer. (G) Heater supply leads. (H) Port for the substrate. (K) Spring clamp. (L) Substrate mask and holder.

temperature of 250°C.

The other substrate holder arrangement was in the form of a large disc with several ports along its rim. It was masked by a fixed disc mounted underneath, that carried a window. The disc with the ports could be turned by mechanical means to position any one of the ports against the window for evaporation. The ports carried a few glass substrates and also the best electropolished Si slice in one of the ports. Fig. (13) illustrates these details of the experimental set-up of the vacuum evaporation technique.

It was desired to shield the substrates during the first portion of the evaporation cycle, since, during this portion the rates of evaporation of In and Sb differed so that the film deposited would be more rich in antimony. In the case of the movable substrate holder, shielding was already provided since the disc with the window shielded all the ports except the one against the window. However, in the case of fixed substrate holder, an aluminum foil covering the substrate mask, served the purpose of the shield. This shield was attached by a piece of string to a small piece of steel screw, which rested inside on the rim of the well close to the bell jar. By manipulation with a magnet the steel piece could be moved, thereby pulling out the shield of the fixed substrate holder.

The heating of the substrates was started along with the outgassing operation of the vacuum system. A rough pump reduced the pressure to 10^{-5} mm. The Ultek ion pump was then started to complete the evacuation, reducing the pressure to 5×10^{-9} mm of Hg. The evacuation took nearly two hours, by which time the substrate temperature had stabilized at 250° C.

The heating current was then turned on and the evaporation cycle was started. The fixed substrate shield was held in place during the first

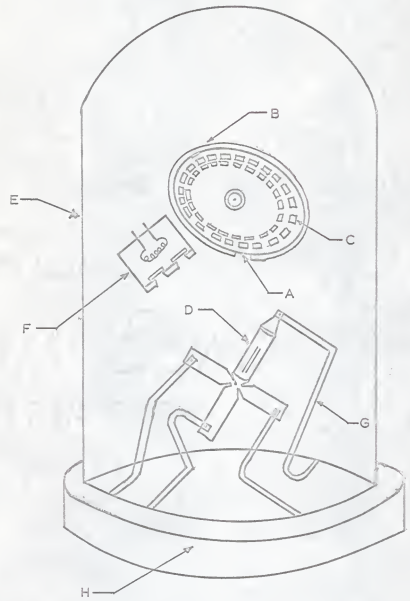


Fig. 13. Vacuum evaporation system. (A) Moveable substrate holder. (B) Mask. (C) Substrate port. (D) Crucible of InSb(shown enlarged). (E) Bell jar. (F) Fixed substrate holder. (G) Current electrode. (H) Vacuum system well.

portion of the cycle, lasting an hour. During this period, a few glass substrates on the disc were exposed in sequence for intervals of 20 minutes each. The thickness of the film formed, on the glass substrate helped to estimate the rate of evaporation. The fixed substrate shield was removed an hour after the beginning of the cycle, and the silicon slices were evaporated on for two hours and fifteen minutes. During this second part of the cycle, the electropolished slice on the disc was exposed for about 45 minutes.

Thus after evaporating on all the substrates for sufficient time, the crucible heating current was turned off and the deposited substrates were allowed to cool gradually and left inside the vacuum system for about six hours. The deposited slices were then removed from the vacuum system for device fabrication and for measurements.

Device Fabrication

The InSb coated substrates were divided into two groups: one group was used to fabricate heterodiodes. The second group was used for studying the epitaxy by means of photo micrographs and electron diffraction.

The first step in the fabrication was to nickel plate the InSb coated substrates. The slices were rinsed in distilled water and then nickel plated on both sides by immersion in an electroless Ni plating bath. The plated slice was then broken into chips. A small area on each chip on both sides was masked with wax paint obtained by dissolving wax in trichloroethylene. The wax protected the InSb film on the substrate and similarly the Si on the rear. The chip was then etched for one minute in CP-4 etch containing by volume 3 parts of 48% HF, 5 parts of concentrated

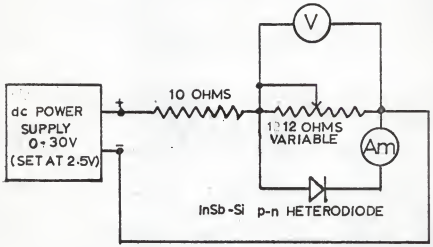
HNO_3, and 3 parts of acetic acid in order to remove the mechanical surface damage introduced on breaking the slice into chips.

After one minute in the etch solution, the etch was diluted away with distilled water to stop the reaction and the chip was removed and rinsed clean in distilled water, then in trichloroethylene to remove the wax mask on its surface. The unetched regions; i.e., those protected by the wax, were used for affixing ohmic contacts. Proper identification of the original chip was made before etching so that the respective sides of InSb and Si were known.

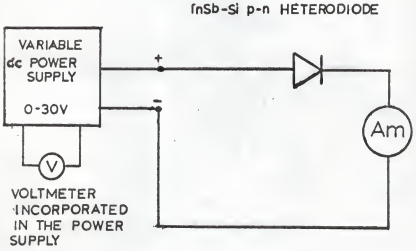
Contacts to Ni plated Si or InSb were made by means of a silver paint thinned with trichloroethylene. Looking through a stereo microscope a thin wire of copper was laid on nickel plating on the InSb side and a globule of silver paint was placed gently on top of it and it was then allowed to dry. After the contact to the InSb side was completely dry, it was turned over and a similar contact to the Si side was made and it was set aside to dry. The heterodiode thus formed was ready for measurements.

Measurements

First of all, the type of impurity present on the deposited film was determined, by testing one of the glass substrates with InSb film on it deposited during the same cycle of evaporation. This test was done using a thermal probe method in which polarity of the hot junction is determined with a galvanometer. Since the majority carriers diffused away from the hot junction, by detecting the polarity, it is possible to determine the type of the semiconductor. The InSb was found to be p type. As the silicon used was n-type the junction formed should be a p-n heterojunction.



CIRCUIT FOR SMALL CURRENTS . V: VOLTMETER, Am: Ammeter



CIRCUIT FOR LARGE CURRENTS

Fig. 14. Circuits employed to obtain static characteristic curves.

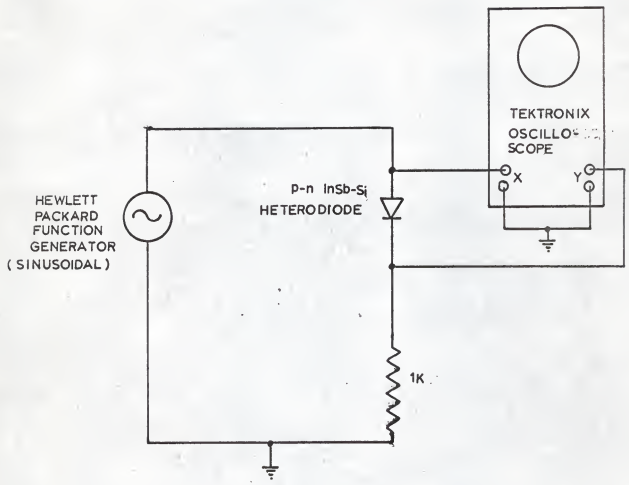


Fig. 15. Schematic of the circuit used to obtain the dynamic characteristic curves.

Measurements were then made using the circuit shown in Fig. (14) for the dc current-voltage characteristic. The dynamic current-voltage curve was found using the circuit shown in Fig. (15), with a Hewlett Packard function generator and a Tektronix oscilloscope.

To study the epitaxy, photo micrographs of the heterojunction were taken with a Leitz high-power microscope. Epitaxial films of InSb on both chemically polished and electropolished silicon were examined. To check the epitaxial nature of the InSb, an electron diffraction pattern was obtained using reflection electron microscopy with the RCA electron microscope.

EXPERIMENTAL RESULTS AND DISCUSSION

Epitaxy

The oriented growth of a crystalline substance on the substrate of another crystalline substance is referred to as epitaxy. The study of epitaxial growth of InSb on silicon single crystal surfaces has been part of this investigation. InSb has a lattice constant of 6.48Å, while silicon has a lattice constant of 5.4Å. The misfit between these two lattice constants is 20%. Pashley, (21) in his study of epitaxy came to the conclusion that small misfit (typically 4% or less) between lattice constants of substances involved is not an essential criterion for epitaxy. In some cases indeed, epitaxial growth has been found to occur with a lattice mismatch of 30%. Thus the mismatch in the case of InSb-Si is not considered to be a deterrent for epitaxial growth.

Evidence deduced from a study of the photomicrographs tends to support

the view that the films are epitaxial. Figures (16a, b, and c) are photomicrographs of the surface of the InSb deposit, on a set of three different silicon substrates, the properties and conditions of deposit of which are as follows. In Fig. (16a) the substrate is an n-type silicon of resistivity 20 ohm-cm. This substrate was chemically polished and was heated to a temperature of 250°C for epitaxial deposition.

In Fig. (16b), the substrate is an n-type silicon of resistivity 20 ohm-cm. This substrate was the best electropolished one with few etch pits and a very flat and smooth surface. However this substrate, by virtue of its size (2mm square) could not be mounted on the fixed heated substrate holder carrying the other substrates. This electropolished substrate was positioned along with a few glass substrates on the movable disc arrangement, and was not heated.

In Fig. (16c) the substrate is an n-type silicon of resistivity 100 ohm-cm. This substrate was only partially electropolished; i.e., after electropolishing, the substrate was found to contain several etch pits. This substrate was also situated in the heated substrate holder and was heated to 250°C.

Figure (16a) overlaps a region containing the InSb film and the silicon substrate so that the boundary of the film on the substrate is seen distinctly, and the InSb and silicon regions are easily identified. The InSb film is seen to possess a smooth and uniform texture with white agglommerations distributed evenly throughout the film. Such a fairly smooth and uniform texture is very characteristic of an epitaxial film as other workers (7, 13, 23) have reported. Glang and Wajda (23) have determined for silicon epitaxy that such a deposit, which may be described as 'micro orange peel', defines a perfect crystal film. Whereas it is

Fig. 16a.

Fig. 16b.

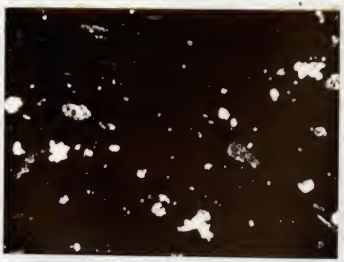
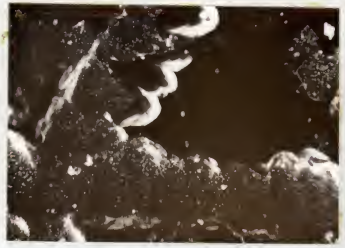


Fig. 16c.



Figs. 16a,b and c. Photomicrographs of InSb films grown on silicon substrates.

difficult to claim with absolute certainty that the film in question here is indeed epitaxial without extensive X-ray analysis, it can be established that the film here is not polycrystalline since polycrystalline films have been found to be in general quite non-uniform and far from being smooth. Further, the arguments presented earlier lead one to believe that the deposit is epitaxial and probably a perfect single crystal film. In contrast to this, the electropolished, but unheated substrate of Fig. (16b) did not develop a smooth epitaxial film, but only scattered agglomerations of InSb crystallites. It has been established that for silicon epitaxy the substrate should be heated. Thus it seems plausible, that the electropolished substrate did not have a smooth epitaxial film since it was not heated during the deposition. However it seems reasonable to surmise that the agglomerations could be single crystalline in nature. The electron diffraction pattern obtained, shown in Fig. (17), tends to support this view. The electron diffraction pattern was obtained by means of reflection electron microscopy. The fact that a spot pattern rather than rings was obtained, indicated the possibility, that the deposit should contain at least single crystallites even if the entire film is not single crystal. Thus the single crystal agglomerates are at least as big as the diameter of the electron beam. Further, rotating the film on a horizontal axis did not change the orientation of the pattern indicating the single crystal nature of the agglomeration.

Finally it is interesting to compare the above photomicrograph with Fig. (16c). The substrate in question here, did not electropolish to the same degree and there were several etch pits on its surface. The faceted nature of the growth seen in the figure may be explained as due to the presence of the etch pits. Other workers [7, 13] have observed such faceted growth on etched substrates. The lack of uniformity and the

Fig. 17. An electron diffraction pattern of the InSb film grown on electropolished silicon substrate.

faceted growth together seem to establish that the deposit is not single crystalline, even in the agglomerations, which implies that it is polycrystalline. Thus a highly polished smooth surface, with the absence of any etch pits, seems to be the ideal surface required for epitaxy.

Current-Voltage Characteristics

The static current voltage characteristics obtained experimentally are indicated in Figs. (18a, b, and c). To bias to the low resistance state it was necessary to bias the InSb positive with respect to the Si. This is in agreement with the energy band diagram given earlier in Fig. (7). The dynamic characteristic curves obtained by means of an oscilloscope trace are shown in Figs. (19a, b, and c).

The static characteristics of a p-n heterojunction were similar to those reported for a homojunction, in general. The forward current varies approximately exponentially with applied voltage and the reverse characteristics show a soft break down. The current-voltage characteristic of p-n junctions including a heterojunction has been derived earlier to be of the form

$$J = J_o \left\{ \exp\left(\frac{qV}{KT}\right) - 1 \right\} \tag{38}$$

However this ideal characteristic is seldom found in total agreement with the one found in practical devices. Thus in order to take into account, the disagreement between theory and experiment, a factor η is introduced and the equation is rewritten as

$$J = J_o \left\{ \exp\left(\frac{qV}{\eta KT}\right) - 1 \right\} \tag{39}$$

The empirical factor η is usually greater than one and approaches the theoretical value of unity only in Ge p-n junctions, and found to be

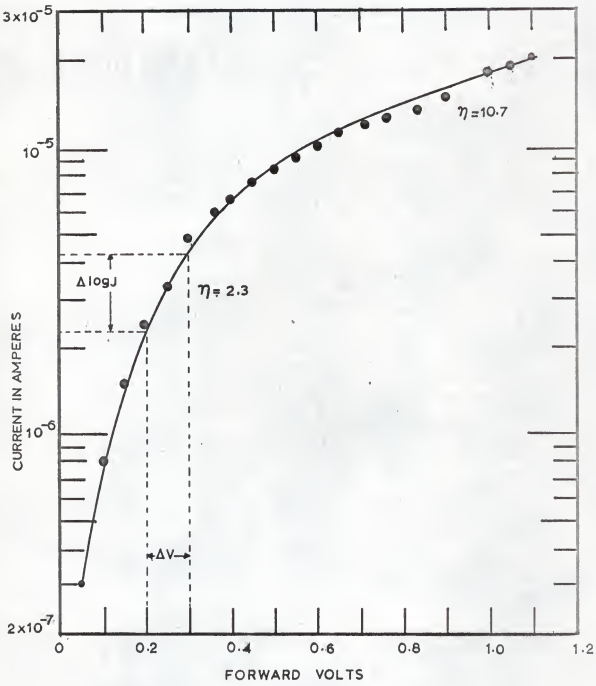


Fig. 18a. Semilogarithmic current-voltage characteristics (log J vs V) of a forward biased p-n InSb-Si heterojunction.

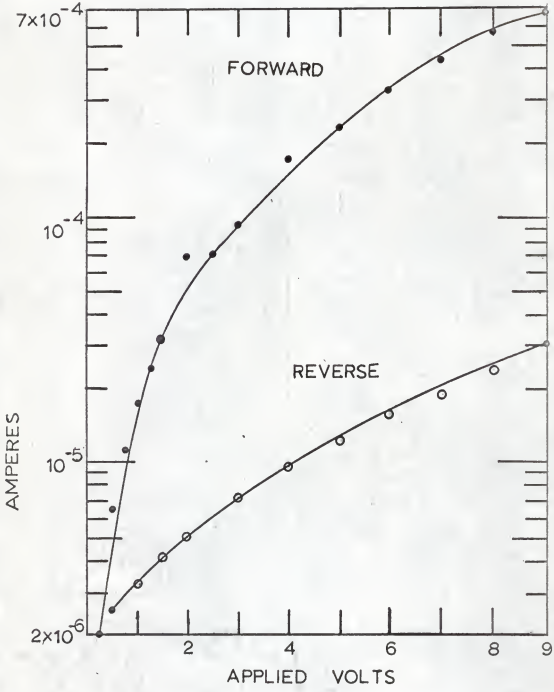


Fig. 18b. Forward and reverse log J vs V characteristics of the p-n InSb-Si heterojunction for large currents.

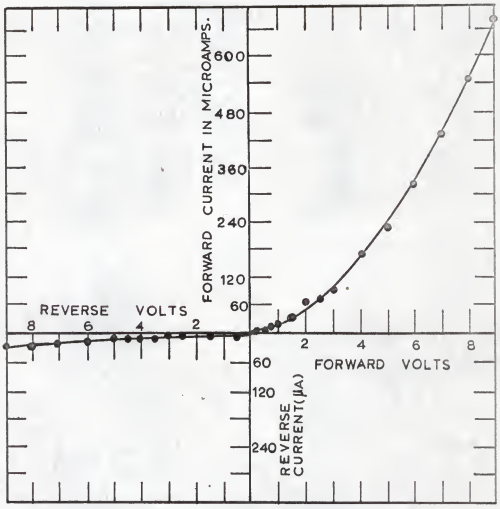


Fig. 18c. Linear current vs voltage plot or the rectifier characteristic of the p-n InSb-Si heterodiode.

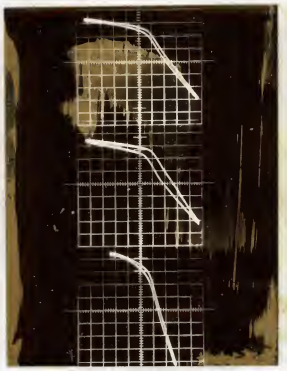


Fig. 19a.
Vertical : 5 mV/div.
Horizontal : 0.5 V/div.
Frequency : 60 cps.
Relative input amplitude : 20

Fig. 19b.
Vertical : 5 mV/div.
Horizontal : 0.5 V/div.
Frequency : 400 cps.
Relative input amplitude : 20

Fig. 19c.
Vertical : 2 mV/div.
Horizontal : 0.5 V/div.
Frequency : 60 cps.
Relative input amplitude : 15

Fig. 19a,b and c. Oscilloscope traces of the dynamic characteristic curves of the p-n InSb-Si heterojunction. Magnitudes of the horizontal and vertical scales, signal frequency and relative input signal amplitude are indicated above.

between 1 and 2 for n-n and p-p heterojunctions of Ge-GaAs. η of the experimental curves was found from the slope of the forward current-voltage characteristic plotted on a semilogarithmic paper shown in Fig. (18a). The ordinate is logarithmic J and the abcissa is the forward voltage in the linear scale. We may obtain from Eq. (39) by manipulation the slope

$$\frac{d}{dv} \log J = \frac{q}{\eta kT} \tag{40}$$

$$\text{or} \quad \Delta \log J = \frac{q}{\eta KT} \Delta v. \tag{41}$$

Using the above equations values of η fitting the small current characteristic of Fig. (18a) were found. η is found to vary and the two values found were $\eta = 2.3$ for very low currents and $\eta = 10.7$ for larger currents.

A similar characteristic has been reported by Anderson[8] for a p-n heterojunction of Ge-GaAs, is also shown in Fig. (20). There is a general agreement as regards the shape of the characteristic between the p-n InSb-Si heterojunction discussed here and the p-n Ge-GaAs reported by Anderson. Anderson reported a sloppy characteristic for the p-n heterojunctions. This is found to be true in the InSb-Si system also. However Anderson obtained much larger current. This primarily depends on the junction area and the relative dopings of the semiconductors. In general, current carriers in heterojunctions are either holes or electrons. In a p-n heterojunction charge is transported by electrons. Since electrons are the majority carriers in the n-type silicon, and since the current is due to electrons injected across the barrier to the p type InSb, the junction current density is proportional to the carrier concentration in the silicon and is thus directly proportional to the conductivity of the silicon or inversely proportional to its resistivity.

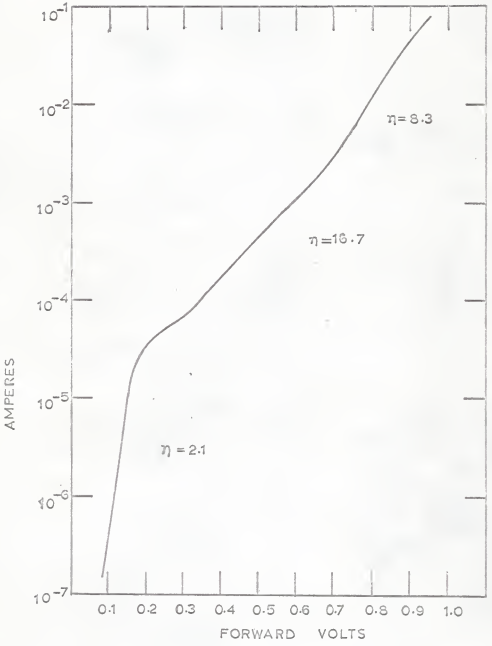


Fig. 20. Semilog plot of current vs voltage characteristics of a forward biased p-n Ge-GaAs (germanium-gallium arsenide) heterojunction at 78° K (after Anderson (8)).

In the device fabricated the resistivities of silicon used were greater than 20 ohm-cm. So the carrier concentrations in the p-n heterojunction of InSb-Si would be very small in comparison to Anderson's p-n Ge-GaAs heterodiodes. This explanation seems to account for the low current densities observed.

Anderson accounted for the different values of η he found in regions a, b, and c (see Fig. 20) by explaining that in region a the value of is small, since the entire applied voltage is effective in altering the barrier to electrons. In region b, only that portion of the applied voltage which appears on the GaAs side is effective in lowering the barrier. Since the resulting current density is due to only a part of the total applied voltage, it will be smaller in comparison to the value given by the ideal rectifier characteristic of Eq. (38) with $\eta = 1$, but will however, be given by Eq. (39) if an increased value of η is taken into account. This is illustrated in Fig. (2). A similar explanation for the log J vs. V characteristics of InSb-Si heterojunction seems to hold well in explaining the different values of η observed (see Fig. 18a).

The large-current characteristics of log J vs. V for both forward and reverse currents were plotted in Fig. (18b). One noticeable feature is that the reverse current did not saturate. Anderson also found this to be the case for p-n Ge-GaAs heterojunction. The reason for this is not known.

Anderson further encountered smaller currents in heterojunction than predicted by theory, for which he gave the following explanation:

> This small value of current might be accounted for by the large reflection expected at the interface, as a result of matching incident, transmitted and reflected Bloch waves. The Bloch waves would be expected to be radically different in the two materials because of the difference in band structure and potential energy on either side of the interface. (Anderson[6], 1960).

It seems probable that a similar explanation could be extended to the InSb-Si heterojunction to account for small order of currents encountered. It is also believed that the smaller current is partly due to the fact that the silicon used was of high resistivity n-type greater than 20 ohm-cm. Figure (19a, b) given the oscilloscope trace of the heterojunction for 60 cps and 400 cps. The trace was obtained by employing circuit shown in Fig. (15), as explained before. A "hysterisis loop" appears as the signal frequency is increased from 60 cps to 400 cps. Apparently a phase shift seems to be introduced at larger frequencies. One possible reason is an increase in the barrier capacitance resulting in the phase shift. However this is in conflict with the theory of heterojunctions since it is known that high injection efficiency of the heterojunction effectively lowers the transition capacitance. The exact cause of the phase shift is not known. Figure (19) gives the oscilloscope trace of the characteristic at 60c/s for a smaller signal. It is seen that the shape of the dynamic characteristic in general, agrees with the static curves.

CONCLUSIONS

Electropolishing

Partial results were obtained with electropolishing and it is believed that the instrument constructed for this purpose can be improved to yield better results by (i) using a cathode with the required perforations, yet with a more perfectly flat surface, (ii) designing a better anode block to facilitate easier planar mounting of the anode slices, (iii) using a better wear-resistant nylon or cotton separator as thin as practicable (preferably

50 to 75 microns) and (iv) using an improved flow system to inject the electrolyte close to the anode slices.

In spite of the limitations of the instrument constructed, excellent polishing was obtained with the 20 ohm-cm, n-type silicon; however, rather poor results were obtained with the 100 ohm-cm, n-type silicon. It is suggested that this is principally due to the high resistivity of the latter coupled with the other defects and limitations in the instrument, improvements for which have been mentioned above.

InSb-Si p-n Heterojunction

In conclusion the InSb-Si p-n heterojunction appears to have properties very similar to the Ge-GaAs p-n heterojunction reported by Anderson. Both the junctions have large differences between the band gaps (1.4-.66 = .74 eV for Ge-GaAs; 1.08-.17 = .91 eV for InSb-Si). Although the lattice mismatch in the InSb-Si heterojunction (about 20%) is much larger compared to the Ge-GaAs system (less than 1%), the characteristics were found to be similar, except in the order of magnitude of the current. This smaller current in InSb-Si can be accounted for by the relatively small doping of the silicon used and by the large carrier reflection at the interface. Thus it appears that the discontinuity in the band edges is more important than the lattice mismatch in governing the characteristics, although it is difficult to say exactly to what extent the lattice mismatch affects the electrical characteristics. The mismatch is believed to increase the denstiy of dislocations and dangling bounds etc., causing interface states [10].

Although the InSb-Si system has much larger mismatch (20%) than the Ge-GaAs (1%), the current-voltage characteristic of the former is no more "sloppy" or "irregular" than the latter. The words "sloppy" or "irregular"

are used in the sense to indicate the departure from the ideal. Thus, the current-voltage characteristics seem to be independent of the mismatch. Further, p-n heterojunctions on both chemically polished and electropolished substrates did not differ in their electrical characteristics. It is expected that any improvement due to electropolishing will be lost in the inherent "irregularity" of the p-n heterojunction characteristic. Though the response to pulse input of a p-n, InSb-Si heterojunction has not been investigated here, it is expected that fast switching performance will result. This is because in this p-n heterojunction conduction is primarily by electrons, the barrier to holes being very large. Thus minority carrier storage effects are avoided when the junction is switched from forward conduction state to reverse bias.

Summary

In summary, the p-n InSb-Si heterojunction characteristics correspond to the energy band diagram derived on the basis of the Anderson model (Fig. 7). Further, the experimental current-voltage curves obtained were similar to those reported by Anderson for another p-n heterojunction system of Ge-GaAs. The discontinuities in band edges seem to affect the characteristics appreciably. It seems probable that the lattice mismatch between InSb and Si did not affect either the epitaxy, or the electrical characteristic.

ACKNOWLEDGEMENTS

The author wishes to express his sincere gratitude to his major professor and research advisor, Dr. E. Brock Dale, for his guidance, encouragement, and keen interest in this project.

Appreciation is expressed to Mr. Grant Poole and Mr. Gerald Hart for their helpful assistance, and to the authorities of the Department of Physics for financial assistance. Appreciation is also expressed to Mr. Anthony Bonnetti of the Sigmund Cohn Corporation for his assistance in goldplating the cathode of the electropolishing mechanism.

BIBLIOGRAPHY

1. Middlebrook. R. A., An Introduction to Junction Transistor Theory, (John Wiley and Sons, Inc., New York, 1957), Chapters 4-7.

2. Van Der Ziel, A., Solid State Physical Electronics, (Prentice Hall Inc., Englewood Cliffs, N. J., 1975), Chapters 12 and 13.

3. Shockley, William, Theory of p-n Junctions in Semiconductors and p-n Junction Transistors. Bell System, Tech. J. 28, 435 (1949).

4. Kroemer, H., Theory of a Wide Gap Emitter for Transistors, Proc. I.R.E. 45, 1535 (1957).

5. Shockley, W., U. S. Pat. 2,569,347.

6. Anderson, R. L., Germanium-Gallium Arsenide Heterojunctions, IBM J. Res. Dev., 4, 283 (1960).

7. Weinstein et al, M., Preparation and Properties of GaAs-GaP, GaAs-Ge and GaP-Ge Heterojunctions, J. Electrochem. Soc., 111, 674 (1964).

8. Anderson, R. L., Experiments on Ge-GaAs Heterojunctions, Solid State Electron., 5, 341 (1962).

9. Oldham, W. G., and Milnes, A. G., n-n Semiconductor Heterojunctions, Solid State Electron., 6, 121 (1963).

10. Oldham, W. G., and Milnes, A. G., Interface States in Abrupt Semiconductor Heterojunctions, Solid State Electron., 7, 153 (1964).

11. Perlman, S. S., and Feucht, D. L., p-n Heterojunctions, Solid State Electron., 7, 911 (1964).

12. Nathan, Marshall I., and Marinace, J. C., Phonon and Polaron Interaction in Germanium-Gallium Arsenide Tunnel Heterojunctions, Phys. Rev., 128, 2149 (1962).

13. Aven, M., and Garawacki, W., Epitaxial Growth and Properties of ZnTe-CdS Heterojunctions, J. Electrochem. Soc., 110, 412 (1963).

14. Sullivan et al, M. V., An Electropolishing Technique for Germanium and Silicon, J. Electrochem. Soc., 110 412 (1963).

15. Baker, D., Tillman, J. R., The Preparation of Very Flat Surfaces of Silicon by Electropolishing, Solid State Electron., 6, 589 (1963).

16. Turner, D. R., Experimental Information on Electrochemical Reaction at Germanium and Silicon Surfaces, in "Electrochemistry of Semiconductors", P. J. Holmes, Editor, (Academic Press Inc., New York, 1962), pp. 155-204.

17. Irving, B. A., Chemical Etching of Semiconductors, in "Electrochemistry of Semiconductors", P. J. Homes, Editor, (Academic Press Inc., New York, 1962), pp. 256-289.

18. Sullivan, M. V., and Warner, R. M., Ohmic Contacts to Silicon and germanium in "Transistor Technology, 3, F. J. Biondi, Editor (D. Van Nostrand Company, Inc., Princeton, N. J., 1958), p. 173.

19. Holmes, P. J., Practical Applications of Chemical Etching, in "Electrochemistry of Semiconductors", P. J. Holmes, Editor (Academic Press, Inc., New York 1962), pp. 329-377.

20. Dale, E. Brock, and Chai, An-Ti, Single Source Technique for the Vacuum Deposition of Alloy Films, Bull, Am. Phys. Soc. 10, 364 (1965).

21. Pahsley, D. W., The Study of Epitaxy in Thin Surface Films, Advances in Physics, Supplement to the Philosophical Magazine, 5-6, 173 (1965).

22. Hulme, K. I., and Mullin, J. B., Indium Antimonide--A Review of Its Preparation, Properties and Device Applications, Solid State Electron., 5, 212-214 (1962).

23. Glang, R., and Wajda, E. S., "Matallurgy of Semiconductor Materials", J. B. Schroeder, Editorm, Metallurgical Society Conference, (Interscience, New York, 1961), p. 27.

24. Kroemer, H., Proposed Class of Heterojunction Injection Lasers, Proc. I.E.E., 51, 1782 (1963).

25. Klein, D. L., and Stefan, D. J., Controlled Etching of Silicon in the $HF-HNO_3$ System, J. Electrochem. Soc., 109, 37 (1962).

p-n HETEROJUNCTIONS OF INDIUM ANTIMONIDE ON ELECTROPOLISHED AND CHEMICALLY POLISHED SILICON

by

RAMASWAMI SUBRAMANIAN

B. Sc., University of Madras, 1956

AN ABSTRACT OF A MASTER'S THESIS

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Physics

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1966

ABSTRACT

A p-n heterojunction of InSb-Si was prepared by epitaxial vacuum evaporation technique. The silicon substrates for this purpose were required to be highly polished, and extremely flat. To fulfill this requirement an electropolishing mechanism was built employing the principle of anodic dissolution and intensive stirring through a system of rotating electrodes. In addition, a chemical polishing method was also employed in the preparation of substrate surfaces.

The electropolishing technique yielded partial results. An extremely good polish was obtained with 20 ohm-cm n-type silicon. The results with 100 ohm-cm n-type silicon were rather poor. Possible improvements to the electropolishing mechanism are suggested.

The films of indium antimonide (InSb) on a group of silicon (Si) substrates with electropolished as well as chemically polished surfaces were examined for epitaxy. From qualitative considerations it is believed that some of the films were of epitaxial growth.

Heterodiodes of InSb-Si were fabricated and the p-n heterojunction thus developed was analyzed for dc and dynamic characteristics. The current-voltage characteristic of the p-n InSb-Si heterojunction obtained, departed from the ideal rectifier characteristic and could be explained by means of the Eq. $J = J_o\left[\exp\left(qv/\eta kT\right) - 1\right]$ where η has values ranging from 2.3 to 10.7. These results were similar to the results obtained by Anderson for Ge-GaAs p-n heterojunction.

An energy band diagram for the InSb-Si system is given corresponding to the Anderson model. On the basis of such a model an attempt is made to explain some of the observed results of the p-n InSb-Si heterojunction.